International Journal of Computer (IJC)

ISSN 2307-4523 (Print & Online)

http://ijcjournal.org/

Analysis of Saudi Arabian Social Network Using Analytic Measures and Community Detection

Lulwah AlSuwaidan[a]*, Mourad Ykhlef[b]

[a]*Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia, Information Management Department, College of Computer and Information Sciences, Al-Imam Mohammad bin Saud Islamic University, Riyadh, Saudi Arabia*

[b]*Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia*

[a]*Email: lnsuwaidan@imamu.edu.sa*, [b]*Email: ykhlef@ksu.edu.sa*

Abstract

Recently, Social Network Analysis has received an enormous popularity in the field of social and computer sciences. The majority of the studied problems have concentrated on research of information diffusion and social influence. The aim of this research is to analyze the Saudi Arabian social network to measure its capability for information diffusion. We are targeting Saudi Arabian social network because of its importance within Arab region. It is considered the most dominant and influence among the others. Social Network Analysis measures (degree, closeness, betweenness, and eigenvector). Community detection, on the other hand, has guaranteed its ability in identifying corresponding community depends on social properties, network structure, or influencers interests. In this article, Griven-Newman community detection algorithm has been adopted to identify the corresponding community. It has been tested and visualized using NodeXL tool. Experiment was applied on Twitter users. The communities resulted and analysis measures' results showed the suitability of the Saudi Arabian network for information diffusion.

Keywords: Social network analysis; information diffusion; social influence; social media.

1. Introduction

Despite the recent joint of Saudi Arabia into social media life, it becomes one of the most influential social community. Benjamin Ampen, head of sales, Middle East and North Africa at Twitter and Marwan Zein of LinkedIn pointed out to the hugeness of the number of users and their activities in Twitter and LinkedIn.

----

* Corresponding author.

This was during their speech at the Workshop on Future of Social Media organized by Saudi Aramco[1]. According to the Economist article on March, 2015, the Saudi Arabia has the world's highest percentage of people on Twitter relative to its number of internet users; and on YouTube too. Saudis also spend more hours online than their peers elsewhere[2]. All the aforementioned statistics put more attention in studying Saudi Arabian network in order to adapt it for many domains especially in business and marketing. In this article, a deep investigation and analysis of this network will be used to measure how important is the network for the problem of information diffusion [3].

Social Network Analysis (SNA) is a mathematical study of vertices and edges that forming the social network [4, 5]. It has a collection of measures used to examine social relationship, interaction among actors, identifying subgraphs, and strong or weak ties [6]. The need of SNA has increased recently because of its effectiveness in various domains including education, business, information sciences, social sciences, information diffusion, and marketing. Studying how nodes in a graph is related and interconnected is highly recommended in hitting the right community using community detection methods.

Community detection, on the other hand, has the property of classifying the network into related groups, each of which has its characteristics depending on the problem needs. Many researches have considered it as a crucial aspect of SNA [7, 8]. The attempts have increased recently in enhancing community detection efficiency and effectiveness especially for large social networks [9].

In this article, the motivation is on analyzing the ability of nodes in spreading the information through the network. In addition, it aims to understand how network structure and vertices' relationships can influence the spreading process. The analysis is done on a sample of Saudi Arabian Twitter network. It is collected at particular time and then analyze it using SNA measures and community detection. It is extracted using NodeXL tool.

The rest of the article is organized as follows: section 2 describes the related works. Then, section 3 presents a discussion of SNA measures. Next, section 4 provides set of dominant community detection methods. The results of the experiments are given in section 5. Finally, the conclusion is presented in section 6.

2. Related works

Social network can be shown as a set of nodes and edges forming a network or graph where nodes are the participants and edges are the type of connections. It has been defined by many studies, the majority have concentrated on representing social network as a graph consisting of nodes and edges. Easley & Kleinberg [10] definition is "A graph specifies relationships among a collection of items.

A graph consists of a set of objects, called nodes, with certain pairs of these objects connected by links called edges", and Kempe and his colleagues [11] definition is "the graph of relationship and interactions within a group of individuals". Social network graph structure is illustrated in Figure 1 which demonstrates the general case (unweighted and undirected).

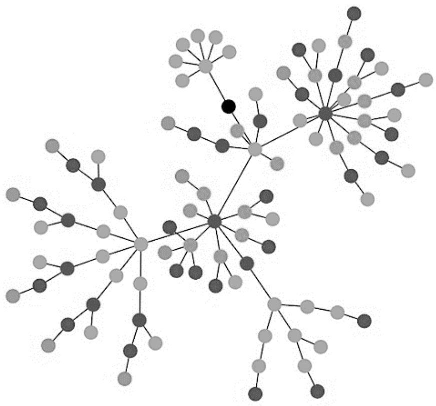

Figure1: Social network graph

A recent survey presented by [6] that discussed a SNA for Counter-Terrorism. They claim the effectiveness of SNA measures and tools in detecting the key players of certain community. They also discussed the capability of SNA methods in discovering and detecting community behavior. Park, Lim, and Park [12] did a comparison between Twitter and YouTube network in terms of information diffusion. They studied both networks from different prospective: network structure, interaction pattern, and geographic distribution of users involved in communication networks of the Occupy Wall Street movement. They found Twitter more suitable for exchanging and facilitating the information while YouTube is more appropriate for diffusing the ideas between different groups. Reference [13] presented a SNA of official Twitter accounts activated during the Charleston, West Virginia, water contamination crisis in 2014. The network had a 41 Twitter accounts associated with the West Virginia water contamination lacked density, contained several isolates, exchanged information quickly (geodesic distance diameter), and contained both national and local accounts. This kind of social analysis assists in determining the limitation and strength in the network which assists for further user of information dissemination and exchange.

Measuring node influence is critical step in influencing process. SNA plays a demanded role in this process by using its measures. Various studies have contributed in this domain [14-17]. Reference [14] proposed a back-propagation neural network to find the user's attributes. This have used to measure the level of users' influence in ResearchGate network. Maharani and his colleagues [15] used degree centrality and eigenvector centrality to find the influencers for marketing in social network. Generally, SNA measures categorized into four main metrics: degree centrality, distance/closeness centrality, betweenness centrality, and eigenvector centrality. All types will be discussed in the next section.

Community detection is one of the most widely used methods in the literature on SNA. It relies on the idea of targeting the communities that are most closely related to a certain domain through social networks. Many researchers have raised questions about the definition of community and its structure. Reference [18] highlighted the need for a clear definition of community; he claimed that this definition is dependent on its context and applications. Reference [19,18] defined community as a group of nodes that are densely connected with each other more frequently than with those outside the group. However, Most of existing methods perform community detection at random basis [20]. The focus of community detection is on the community structure because of its usefulness in knowing how the network is structured [21]. Recently, the interest has changed to cover both topological and topical community detection [7]. Reference [22] introduced a community detection approach to viral marketing. Their work concentrated on identifying overlapping communities in online social networks. Two measures, node betweenness and probability, were used to measure the overlapping influence for nodes. If the probability lay between 1% and 5%, then it became a top influencer node. This supported the hypothesis proposed by [22]: the more a community overlaps, the greater the individual's influence in the whole network.

3. Social network analysis measures

Reference [23] proposed three different intuitive notion of centrality namely degree, betweenness and closeness centrality, which are mostly used in identifying key players in the social network [24]. In this section, details in each of these measures and eigenvector centrality will be discussed.

3.1. Degree centrality

Degree centrality measures the node importance which is computed by the number of adjacent nodes. The larger number of connected nodes reflects the high importance of nodes [25, 26]. Considering the high-degree nodes as influential in social networks is the key element in degree centrality. Node degree is the degree centrality, so knowing the node degree inferences the node centrality degree. This metric categorized by Alahakoon and his colleagues [27] into local and global centrality metrics. Local metrics refer to local centrality that connect local nodes reside within one community structure. On the other hand, global centrality refers to centrality betweenness that connects two nodes in two different community structures. The degree centrality computed from the following Eq. (1):

$$C_D(v_i) = d_i = \sum_d A_{ij} \tag{1}$$

where d_i is the number of node i neighbours in a network.

In order to measure the degree centrality between two nodes in two different networks, a normalized degree centrality has to be computed as Eq. (2).

$$C_d^i(V_i) = \frac{d_i}{n-1} \tag{2}$$

Moreover, degree centrality measures how vertices are communicating within community. The more central vertices are more related to community than leaf ones. In addition, node location plays an important role in communities of social network. Nodes might be bridge/hub to connect two or more communities while outlier nodes only connect by single line to one community [28].

3.2. Distance/closeness centrality

This measure concerns about the shortest path between node and other nodes in the network. It considered as an influence node measure since the node with short path with other nodes will have higher influence than any other node [11]. Closeness is slightly different than distance centrality. It also takes into consideration the short paths but by averaging all shortest paths of node to all other nodes[25]. It can be computed by using the following equation (3) [26]:

$$C_c(v_i) = \frac{n-1}{\sum_{j\neq i}^{n} g(v_i, v_j)} \quad (3)$$

where n is the number of nodes, and $g(v_i, v_j)$denotes geodesics distance between node v_i and v_j.

3.3. Betweenness centrality

Node betweenness defined by [26] as the number of shortest paths that pass one node in the network. Once the node has high betweenness centrality, it would have a better communication and diffusion. The shortest paths between nodes in different clusters have to go through the few interclusters connections, which therefore have a large betweenness value [29]. It computed by the following Eq. (4):

$$C_B(v_i) = \sum_{v_s \neq v_i \neq v_t \in V, s<t} \frac{\sigma_{st}(v_i)}{\sigma_{st}} \quad (4)$$

where σ_{st} is the number of shortest paths between t and s, $\sigma_{st}(v_i)$ is the number of shortest paths between t and s via v_i.

The betweenness centrality has been defined by [25] as the following Eq. (5):

$$C_i^{BET} = \sum_{j,k} \frac{b_{ijk}}{b_{jk}} \quad (5)$$

where b_{jk} is the number of shortest paths from node j to k, b_{ijk} is the number of shortest paths from node j to k that pass through node i.

In addition, a random walk based betweenness centrality measure proposed by [30]. It has been applied on the graph structure. The idea behind this alternative measure of betweenness centrality is to compute the betweenness from different random walks. The Newman's betweenness centrality C_i^{BET} of node i can be measured by using the following Eq. (6):

$$C_i^{NBE} = \sum_{j\neq i\neq k} R_{jk}^{(i)} \quad (6)$$

here $R^{(i)}$ be the matrix whose (j, k)th element $R_{jk}^{(i)}$ contains the probability of a random walk from j to k, which contains i as an intermediate node.

There are many betweenness centrality measures have been discussed in the literature. First, K-path centrality measure which counts the number of paths. The k metric in k-path method refers to the largest length connecting two nodes. This method was proposed by Alahakoon and his colleagues [27]. It based on betweenness centrality score to measure the level of node influence. The algorithm was tested randomly for estimation and was proved that the nodes with high K-path centrality had high betweenness centrality.

3.4. Eigenvector centrality

This centrality measure aims at measuring the node importance within a network based on the importance of its neighbours. There are many centrality measures were proposed in the literature that cover this kind of measure[11]. First, Katz measure [31] which takes into consideration the network diffusion behavior. The paths starting from a node are counted by Katz measure in purpose of tracking the walks and discarding the longer ones. Jackson [32] defined Katz prestige of node i as $P_i^K(g)$ which corresponds to the sum of prestige of all neighbors divided by the respective degree. Eq. (7) computes Katz prestige of a node i as:

$$P_i^K(g) = \sum_{i \neq j} g_{ij} \frac{P_i^K(g)}{d_j(g)} \quad (7)$$

Since the equation/formula is self-referential, Jackson [32] suggested normalized adjacency matrix as $\hat{g}_{ij} = g_{ij}/d_j(g)$. Therefore, Eq. (7) can be rewritten as shown in Eq. (8):

$$P^k(g) = \hat{g} P^k(g) \quad (8)$$

Some enhancements on Katz measure were done to produce the second prestige measure of Katz. This measure is based on summation of all weighted walks originated from node i. the second Katz prestige measure is shown as Eq. (9):

$$P^{k2}(g, a) = (I - aG)^{-1} aG1 \quad (9)$$

where, I is $n\ x\ n$ identity matrix, 1 is $n\ x\ 1$ vector of 1s, and $0 < a < 1$.

Bonacich centrality [33] is an extension of the second Katz prestige measure. It is given by Eq. (10):

$$C^B(g, a, b) = (I - bG)^{-1} aG1 \quad (10)$$

where $a > 0$ and $b > 0$.

4. Community detection techniques

As mentioned earlier, community detection is considered as a significant tool for analyzing complex networks. There are number of well-known community detection methods that have proven their effectiveness for various domains. The divisive scheme by Girvan and Newman [34], the modularity maximization method of [35] and its computationally efficient implementation [36], the N-cut graph partitioning scheme by [37], and spectral partitioning by use of the graph Laplacian [38]. The contributions in community detection have enormously increased recently. The majority have concentrated on solving problems of large network, and enhancing the performance of detection. Meng and his colleagues [39] have proposed an incremental density-based link clustering algorithm for community detection in dynamic networks called iDBLINK. This algorithm directed to solve the problem in dynamic social network. It ansured its accuracy and efficiency. Probabilistically Mining Communities (PMC) proposed by Yang and his colleagues [9] aimed at optimizing the problem of community detection by use of random walk based heuristic.

5. Experiments and analysis

Measuring social influence was analyzed on Twitter social network. It is an explicit network because users forming their connections by sharing common interest. It was unimodal and extracted using NodeXL[40]. NodeXL allows visualizing through sorting, filtering, and clustering functions. It also provides functions that analyse networks graphs, and computing the network metrics [41].

5.1. Dataset collection

The data used for this study is collected from Twitter. It is a free pool to exchange information within a community. It can be shown as a graph consists of a set of vertices and edges that share a common interest. The common interest might be a topic, a real-world person, a place, an event, an activity or a cause [28]. For instance, the set of all Twitter accounts that tweets, retweets, hashtags and replays to a certain topic such as “Apple Conference”. All accounts participated in this topic will cause a subgraph forming an Apple community. More precisely, Social Media community generation process has classified into explicit or implicit by Papadopoula and his colleagues [28]. Explicit community generation is based on human decisions to form their community. For instance, the Twitter list can be generated by any Twitter account to add members that are interested in certain activity. On the other hand, implicit community generation can be generated implicitly by monitoring the overall members' behaviors. The system is responsible to discover the network and find out members who are sharing the common interest.

The dataset consists of 47739 nodes and 53210 edges. It is concentrated on Saudi Arabia community. The selection of root users was chosen randomly from different domains. The NodeXL started crawling to import the network of each user. The dataset was collected on January 2016.

5.2. Social analysis

As discussed in section 3 and 4, there are different measures for social analysis and community detection. This study has applied all aforementioned measures on the collected dataset. The overall metrics has given a modularity of value 0.86. According to the definition of modularity presented by Leicht and Newman [42], the high modularity value indicates a good quality of division into communities. The density of the graph was given a value 0.000023, which means the graph has limited density. Density should be high in terms of community but lower for the full graph. That indicates a suitability for information diffusion. Diffusion in social network seeks for high modularity in terms of quality division and low density between divisions. That keeps the diffusion go viral within community and lower between them because of limited edges and hubs.

The geodesic distance is the distance between two vertices along the shortest path between them. The average geodesic distance of Saudi Arabia social graph is 3.8, where the maximum geodesic distance is 6. Therefore, only 3 to 6 edges is needed to reach one node.

The analysis measures are indicators of node importance in a graph. NodeXL has computed a various social analysis measures, but this study concentrates on view of them. Since the graph collected is directed, the degree centrality has split into in-degree and out-degree. According to Eq. (1) and Eq. (2), the maximum in-degree is 2006, and maximum out-degree is 2100. Generally, degree centrality measures the popularity of node. Since the degree centrality given is high that means the node is popular within a graph. In terms of betweenness centrality Eq. (4) and Eq. (5), the maximum value given is 355331558.422. As it measures of a node's ability to bridge different subnetworks, the high shows how suitable this node for information diffusion. In addition, eigenvector metric measures the degrees of the nodes that a node is connected to. By using the equations in section 3.4, we obtained the maximum value of 0.008 which is good enough to consider node popularity over its connection.

Social networks has been widely used for the objective of information diffusion. In order to achieve this, the majority of social network analysis studies seek to detect communities in social graph. We have chosen the Griven-Newman algorithm because of many reasons: its ''divisive'' technique which repeatedly removes edges from the network. The removed edges are determined by one of a set of edge betweenness measures. In addition, its recalculation and re-evaluated of betweenness scores along with edge removal process. Also, it is reliably and accurately extract community structure from artificially generated networks with known communities. Figure 2 shows a 21 communities resulted after applying Griven-Newman algorithm on Saudi Arabia dataset. It shows a 21 groups that are dense and overlapped. These properties fit the requirements of information diffusion in network since one piece of information initiated in one community can easily passed to another community.

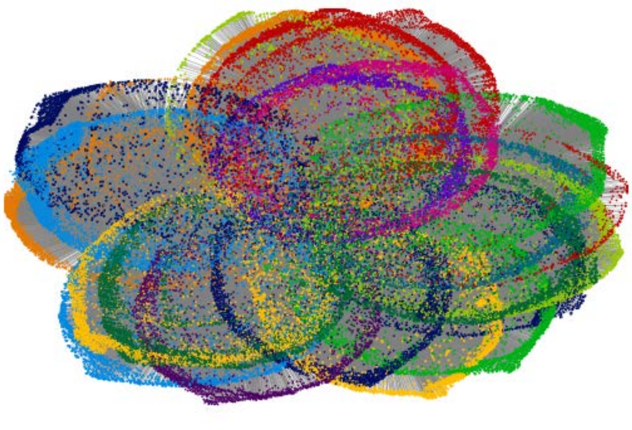

Figure 2: Community network using Girven-Newman community detection algorithm

6. Limitations

Several research studies have used NodeXL for analyzing and visualizing the social networks. Even though this tool is powerful in analysis, it has some limitations that keeps it with low accuracy and consistency. However, NodeXL data does not overcome many of other challenges facing network analysis in general and network visualization in particular. Its layout algorithms usually fail to find optimal arrangements of lines and nodes in order to maximize the comprehension of the structure of the network. In terms of large scale data sets, it remains hard to concisely display [41, 43]. NodeXL is limited to specific period for collecting, for example it allows only the last 7 days to be viewed. It also limits the collecting for the last 1000 tweets, or retweets which affects the accuracy of measurements [44].

7. Conclusion

This work performed a social analysis to measure the capability of network diffusion information throughout its vertices and relationship. The network was extracted and visualized using NodeXL tool. The SNA measures computed were: degree, betweenness, closeness, and eigenvector. These measures have been chosen because of usefulness in identifying the importance of certain node within a network. These measure have given a sufficient values as indicators of suitability for diffusion and influence. Other measures such as modularity, density and geodesic distance. All have given a satisfactory values enough to achieve reasonable diffusion. Lastly, Griven-Newman algorithm has used for community detection, it produced a 21 fair communities with high modularity reached 0.86.

References

[1] "Saudi social media users ranked 7th in world," ed: Arab News, 2015.

[2] "The Economist explains Why Saudis are ardent social media fans," ed: The Economist, 2015.

[3] N. Lenze, "Social Media in the Arab World: Communication and Public Opinion in the Gulf States," European Journal of Communication, vol. 32, pp. 77-79, 2017.

[4] J. A. Barnes and F. Harary, "Graph theory in network analysis," Social Networks, vol. 5, pp. 235-244, 1983.

[5] S. Wasserman and K. Faust, Social network analysis: Methods and applications vol. 8: Cambridge university press, 1994.

[6] P. Choudhary and U. Singh, "A Survey on Social Network Analysis for Counter-Terrorism," International Journal of Computer Applications, vol. 112, 2015.

[7] Z. Zhao, S. Feng, Q. Wang, J. Z. Huang, G. J. Williams, and J. Fan, "Topic oriented community detection through social objects and link analysis in social networks," Knowledge-Based Systems, vol. 26, pp. 164-173, 2012.

[8] S. Fortunato, "Community detection in graphs," Physics Reports, vol. 486, pp. 75-174, 2010.

[9] B. Yang, J. Di, J. Liu, and D. Liu, "Hierarchical community detection with applications to real-world

network analysis," Data & Knowledge Engineering, vol. 83, pp. 20-38, 2013.

[10] D. Easley and J. Kleinberg, Networks,Crowds, and Markets: Reasoning about a Highly Connected World.: Cambridge University Press 2010.

[11] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the Spread of Influence through a Social Network," in KDD '03 Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, New York, 2003, pp. 137-146.

[12] S. J. Park, Y. S. Lim, and H. W. Park, "Comparing Twitter and YouTube networks in information diffusion: The case of the “Occupy Wall Street” movement," Technological Forecasting and Social Change, vol. 95, pp. 208-217, 2015.

[13] M. C. Getchell and T. L. Sellnow, "A network analysis of official Twitter accounts during the West Virginia water crisis," Computers in Human Behavior, vol. 54, pp. 597-606, 2016.

[14] O. Alheyasat, "Investigation and Analysis of Research Gate User’s Activities using Neural Networks," The International Arab Journal of Information Technology, vol. 2, 2016.

[15] W. Maharani, Adiwijaya, and A. A. Gozali, "Degree centrality and eigenvector centrality in twitter," in Telecommunication Systems Services and Applications (TSSA), 2014 8th International Conference on, 2014, pp. 1-5.

[16] U. Brandes, S. P. Borgatti, and L. C. Freeman, "Maintaining the duality of closeness and betweenness centrality," Social Networks, vol. 44, pp. 153-159, 2016.

[17] A. G. Nikolaev, R. Razib, and A. Kucheriya, "On efficient use of entropy centrality for social network analysis and community detection," Social Networks, vol. 40, pp. 154-162, 2015.

[18] H.-W. Shen, Community Structure of Complex Networks: Springer Berlin Heidelberg, 2013.

[19] L. Tang and H. Liu, Community Detection and Mining in Social Media: Morgan & Claypool Publishers, 2010.

[20] P. Zhang, C. Moore, and M. E. J. Newman, "Community detection in networks with unequal groups," Physical review E, vol. 93, p. 012303, 2016.

[21] D. Kempe, Structure and Dynamics of Information in Networks. Los Angeles: University of South California, 2011.

[22] S. Y. Bhat and M. Abulaish, "Overlapping Social Network Communities and Viral Marketing," in In proceeding of: International Symposium on Computational and Business Intelligence, New Delhi, 2013.

[23] L. C. Freeman, "Centrality in social networks conceptual clarification," Social Networks, vol. 1, pp. 215-239, 1978.

[24] L. Alvarez, K. Borsi, and L. Rodrigues, "The role of social network analysis on participation and placemaking," Sustainable Cities and Society, vol. 28, pp. 118-126, 2017.

[25] J. Sun and J. Tang, "A Survey of Models and Algorithms for Social Influence Analysis," in Social Network Data Analytics, C. C. Aggarwal, Ed., ed Boston, MA: Springer US, 2011, pp. 177-214.

[26] L. Tang and H. Liu, "Community detection and mining in social media," Synthesis Lectures on Data Mining and Knowledge Discovery, vol. 2, pp. 1-137, 2010.

[27] T. Alahakoon, R. Tripathi, N. Kourtellis, R. Simha, and A. Iamnitchi, "K-path centrality: a new centrality measure in social networks," presented at the Proceedings of the 4th Workshop on Social Network Systems, Salzburg, Austria, 2011.

[28] S. Papadopoulos, Y. Kompatsiaris, A. Vakali, and P. Spyridonos, "Community detection in social media," Data Mining and Knowledge Discovery, vol. 24, pp. 515-554, 2012.

[29] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 101, pp. 2658-2663, 2004.

[30] M. E. Newman, "A measure of betweenness centrality based on random walks," Social Networks, vol. 27, pp. 39-54, 2005.

[31] L. Katz, "A new status index derived from sociometric analysis," Psychometrika, vol. 18, pp. 39-43, 1953.

[32] M. O. Jackson, Social and economic networks vol. 3: Princeton university press Princeton, 2008.

[33] P. Bonacich, "Power and centrality: A family of measures," American journal of sociology, pp. 1170-1182, 1987.

[34] M. Girvan and M. E. Newman, "Community structure in social and biological networks," Proceedings of the national academy of sciences, vol. 99, pp. 7821-7826, 2002.

[35] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," Physical review E, vol. 69, p. 026113, 2004.

[36] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," Physical review E, vol. 70, p. 066111, 2004.

[37] J. Shi and J. Malik, "Normalized cuts and image segmentation," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, pp. 888-905, 2000.

[38] U. Von Luxburg, "A tutorial on spectral clustering," Statistics and computing, vol. 17, pp. 395-416, 2007.

[39] F. Meng, F. Zhang, M. Zhu, Y. Xing, Z. Wang, and J. Shi, "Incremental Density-Based Link Clustering Algorithm for Community Detection in Dynamic Networks," Mathematical Problems in Engineering, vol. 2016, p. 11, 2016.

[40] D. Hansen, B. Shneiderman, and M. A. Smith, Analyzing social media networks with NodeXL: Insights from a connected world: Morgan Kaufmann, 2010.

[41] M. A. Smith, B. Shneiderman, N. Milic-Frayling, E. M. Rodrigues, V. Barash, C. Dunne, T. Capone, A. Perer, and E. Gleave, "Analyzing (social media) networks with NodeXL," presented at the Proceedings of the fourth international conference on Communities and technologies, University Park, PA, USA, 2009.

[42] E. A. Leicht and M. E. Newman, "Community structure in directed networks," Physical review letters, vol. 100, p. 118703, 2008.

[43] S. Matei, "Analyzing Social Media Networks with NodeXL: Insights from a Connected World by Derek Hansen, Ben Shneiderman, and Marc A. Smith," International Journal of Human–Computer Interaction, vol. 27, pp. 405-408, 2011/02/23 2011.

[44] R. Recuero, R. Araujo, and G. Zago, "How does social capital affect retweets?," in ICWSM, 2011.